**Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

**Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

Team project ©2017
Dony Pratidana S. Hum | Bima Agus Setyawan S. IIP

# BAB II

# LANDASAN TEORI

## 2.1 Grafologi

Grafologi adalah ilmu tulisan tangan yang dihasilkan oleh pikiran atau kerja otak. Berdasarkan hal tersebut, banyak alhi grafologi menyebut tulisan tangan dengan tulisan otak (Rahmi, 2018). Untuk memperoleh informasi tentang kepribadian dapat dilakukan dengan menganalisis unsur-unsur grafis struktural tulisan tangan. Grafologi telah ada sejak zaman dahulu kala, orang-orang dahulu sengaja mempelajari atau menganalis tulisan tangan seseorang demi mencari tahu bagaimana karakternya. Buku pertama yang menjelaskan bagaimana grafologi ditulis oleh Camillo Baldi yaitu seorang dokter asal Itali di tahun 1622.

## 2.2 Karakter Berdasarkan Jenis Tulisan Tangan

Pada penelitian ini diklasifikasikan menjadi 6 kategori tulisan tangan untuk menganalisa karakter beserta penjelasan dari masing-masing kategori.

### 2.2.1 Ukuran Huruf

Ukuran huruf mengindikasikan keinginan untuk diperhatikan. Huruf dibagi menjadi huruf kapital dan huruf non kapital. Huruf kapital ukuran normal memiliki ukuran 9mm dan huruf non kapital normal memiliki ukuran 3mm. Diluar itu dianggap sebagai ukuran huruf besar jika melebihi ukuran huruf normal dan ukuran huruf kecil jika kurang dari ukuran huruf normal (Prasad, dkk., 2010). Pada Tabel 2.1 dijelaskan mengenai karakter berdasarkan ukuran huruf (Taufik, 2010).



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

Tabel 2.1 Karakter Berdasarkan Ukuran Huruf

| No | Kategori | Karakter |
|---|---|---|
| 1 | Besar (>3mm) | Perlu pengakuan orang lain untuk memotivasi diri |
| 2 | Sedang (3mm) | Stabil dan dapat beradaptasi |
| 3 | Kecil (<3mm) | Membatasi diri dalam kehidupan sosial |

### 2.2.2 Kemiringan Tulisan

Kemiringan tulisan menunjukkan bagaimana berinteraksi secara emosional dengan lingkungan sosial (Taufik, 2010). Tulisan dengan kemiringan > $90^0$ merupakan tulsisan dengan kemiringan ke kanan, Bila < $90^0$ merupakan tulisan dengan kemiringan ke kiri. Tulisan vertikal memiliki sudut $90^0$. Pada Tabel 2.2 dijelaskan mengenai Karakter Berdasarkan Kemiringan Kata (Prasad, dkk., 2010).

Tabel 2.2 Karakter Berdasarkan Kemiringan Kata

| No | Kategori | Karakter |
|---|---|---|
| 1 | Vertikal ($90^0$) | Praktikal, mandiri, dapat mengendalikan diri, dapat memenuhi kebutuhan sendiri. |
| 2 | Miring ke kanan (>$90^0$) | Mudah bersosial, mau mendengarkan, tertarik kepada orang lain, ramah. |
| 3 | Miring ke kiri (<$90^0$) | Pendiam, pengamat yang baik, bergantung pada diri sendiri, tidak suka mengganggu atau ikut campur urusan orang lain. |

### 2.2.3 Baseline

*Baseline* menggambarkan suasana hati dan *mood* dalam bekerja. *Baseline* dapat diukur dengan melihat kecenderundan arah huruf apakah naik, turun, lurus atau naik turun dalam satu baris tulisan. Pada Tabel 2.3 dijelaskan mengenai Karakter Berdasarkan *Baseline* (Asra, S. & Shubhangi, D.C., 2015).



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

Tabel 2.3 Karakter Berdasarkan *Baseline*

| No | Kategori | Karakter |
|---|---|---|
| 1 | Naik | Memiliki sikap positif, bersemangat dan aktif. |
| 2 | Lurus | Dapat mengontrol diri, bersemangat dan aktif. |
| 3 | Turun | Memiliki enefti negative, berpikiran negative dan sering kecewa. |

### 2.2.4 Ukuran Margin

Ukuran margin menggambarkan cara mengorganisir diri dan kehidupan. Ukuran sempit adalah kurang dari 1 cm dan ukuran lebar adalah lebih dari 1cm. Pada Tabel 2.4, Tabel 2.5, dan Tabel 2.6 dijelaskan mengenai karakter berdasarkan ukuran margin (Taufik, 2010).

Tabel 2.4 Karakter Berdasarkan Ukuran Margin Atas

| No | Kategori | Karakter |
|---|---|---|
| 1 | Sempit (<1cm) | Sombong |
| 2 | Lebar (>1cm) | Menunjukkan perhatian |

Tabel 2.5 Karakter Berdasarkan Ukuran Margin Kiri

| No | Kategori | Karakter |
|---|---|---|
| 1 | Sempit (<1cm) | Mengingat masa lalu dengan baik |
| 2 | Lebar (>1cm) | Berusaha maju dari masa lalu |

Tabel 2.6 Karakter Berdasarkan Ukuran Margin Kanan

| No | Kategori | Karakter |
|---|---|---|
| 1 | Sempit (>1cm) | Masa lalu tidak membuat takut menghadapi masa depan |
| 2 | Lebar (<1cm) | Masa lalu membuat kurang berani menghadapi masa depan |

UNIVERSITAS MULTIMEDIA NUSANTARA

### 2.2.5 Ukuran Jarak Antar Baris

Ukuran jarak antar baris menggambarkan nyaman atau tidaknya seseorang terhadap orang lain (Rahmi, 2018). Ukuran jarak teratur atau normal antabaris adalah bisa disisipi satu baris tulisan lagi (Taufik, 2010). Pada tabel 2.7 dijelaskan mengenai Karakter Berdasarkan Ukuran Huruf (Amend & Ruiz, 1980).

Tabel 2.7 Karakter Berdasarkan Ukuran Spasi Antarbaris

| No | Kategori | Karakter |
|---|---|---|
| 1 | Kecil | Butuh mengekspresikan diri dalam kata-kata, tindakan, dan rencana. Pemilik tulisan ini cenderung bersemangat, tegas, dan kreatif tetapi kurang memiliki tujuan yang jelas dan kurang konsentrasi. |
| 2 | Lebar | Cenderung mengisolasi diri dari lingkungan sosial, psikologi atau keduanya. Pemilih tulisan ini tumbuh dengan ketakutan akan kontak dan kedekatan dengan orang lain atau memiliki kecurigaan terhadap orang lain yang membuatnya tidak percaya dan menjaga jarak. |
| 3 | Normal | Memiliki harmoni dan fleksibilitas sendiri |

## 2.3 Image Processing

*Image Processing* digunakan untuk meningkatkan kualitas data gambar untuk digunakan pada proses selanjutnya. *Image Processing* pada dasarnya terdiri dari tiga langkah.

1. *Import* gambar dari *optical scanner* atau foto digital
2. Analisis dan manipulasi gambar seperti data *compression*, peningkatan gambar dan pola seperti pada foto satelit
3. *Output* dari gambar yang telah dimodifikasi



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

Pada aplikasi identifikasi karakter berdasarkan tulisan tangan, *image processing* yang dilakukan adalah *binarization*, *dilation*, *contours*, dan *warp affine transformation*.

### 2.3.1 Binarization

Gambar *binary* memiliki value 0 (hitam) atau 255 (putih). Proses *binarization* pada gambar dokumen *binary* dilakukan pada proses *image preprocessing* untuk menganalisis dokumen dan bertujuan untuk memisahkan *foreground text* dari dokumen *background* (Su, dkk, 2013). Untuk membentuk gambar *binary*, *threshold* diaplikasikan sehingga gambar dapat diklasifikasi menjadi *foreground atau background.*

### 2.3.2 Dilation

Proses *dilation* mengubah struktur elemen dan menghasilkan gambar baru. Lapisan pixel akan ditambahkan pada batas dalam dan luar gambar. Hasil dari penggunaan *dilation* pada gambar *gray-scale* adalah jika nilai pada *structuring element* positif, maka gambar *output* cenderung lebih cerah dari pada input. Sedagkan detail gambar yang gelap akan dikurangi atau dieliminasi, tergantung dari bagaimana nilai dan bentuk yang saling berhubungan terhadap *structuring element* pada *dilation* (Gonzales & Woods, 2002).

### 2.3.3 Contours

Setelah proses *dilation*, *contours* gambar dapat ditemukan. *Contours* adalah titik lengkungan pada gambar yang saling berhubungan. Dengan kata lain, *contours* menggambarkan bentuk objek pada gambar. Pendekatan *contour*



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

biasanya dimulai dengan *edge detection*, diikuti dengan proses mencari garis kurva bersambung (Malik, dkk., 2001).

### 2.3.4 Warp Affine Transformation

*Affine Transformation* adalah transformasi yang dapat diekspresikan dalam bentuk *matrix multiplication* diikuti dengan *vector translation*. Transformasi seperti translasi, *scaling* dan rotasi dikategorikan sebagai *affine transformations* karena transformasi tersebut menjaga hubungan linear dan perbandingan jarak. Berarti setiap point yang linear sebelum transformasi akan tetap sama setelah transformasi dan rasio jarak akan mengikuti transformasi (Singala & Goyal, 2012).

## 2.4 Segmentasi

Segmentasi gambar merupakan salah satu pendekatan pada *image processing.* Menurut Gonzales dan Wood (2002), segmentasi gambar merupakan proses membagi-bagi gambar digital dan digunakan untuk mengalokasikan gambar menjadi angka agar lebih mudah dianalisis. Pada proses ini, gambar digital tulisan tangan dibagi menjadi tiga segment yaitu kata, huruf dan garis *baseline.* Tiap segment digunakan untuk menghitung *feature* yang berhubungan dengan ukuran huruf, kemiringan kata, *baseline*, tekanan penulisan, spasi antar kata, ukuran margin atas, margin kiri dan margin kanan. Segmen gambar kemudian akan diproses dengan persamaan matematika yang hasilnya untuk masuk ke dalam proses klasifikasi (Prasad, dkk., 2010).

UNIVERSITAS MULTIMEDIA NUSANTARA

## 2.5 Projeksi Vertikal dan Horizontal

Projeksi vertikal dan horizontal adalah proses *scan* garis secara vertikal dan horizontal. *Linearization* gambar hanya menghasilkan dua jenis warna yaitu hitam dan putih. Jumlah pixel pada *gray value* adalah 0 dan 255 yang akan dikonversi menjadi 1 sebagai *pixel* hitam dan 0 sebagai *pixel* putih. Nilai akan didapat dari *scanning* gambar garis per garis dari atas ke bawah dan didapatkan jumlah pixel pada setiap *scanning* (Zhu, dkk., 2017). Untuk projeksi horizontal dilakukan *scanning* dari atas ke bawah sedangkan untuk vertikal adalah dari kiri ke kanan atau sebaliknya.

## 2.6 Deslanting

Menurut Vinciarelli dan Luettin (2010), kemiringan huruf didapatkan dengan densitas vertikal (jumlah *foreground pixel* per kolom) dibagi dengan jarak antara pixel tertinggi dan pixel terendah dalam satu kolom, dimana kolom adalah huruf dalam satu kata kemudian hasil pembagian tersebut dikuadratkan.

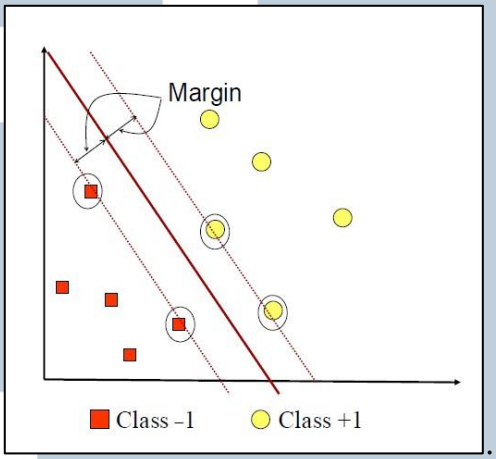
## 2.7 Support Vector Machine

*Support Vector Machine* (SVM) adalah algoritma *supervised machine learning* yang digunakan untuk klasifikasi atau regresi. SVM berusaha menemukan *hyperlane* yang terbaik pada *input space. Hyperlane* adalah pemisah antar dua kelas. *Hyperlane* pemisah terbaik dapat ditemukan dengan memaksimalkan jarak *hyperlane* dengan *pattern* terdekat masing – masing *class.* Jarak *hyperlane* dengan *pattern* terdekat disebut *margin. Pattern* yang paling dekat disebut *support vector* (Nugroho, dkk., 2003).



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

Garis solid pada Gambar 2.1 menunjukkan *hyperlane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*, sedangkan titik merah dan kuning yang berada dalam lingkaran hitam adalah *support vector*.




Gambar 2.1 *Hyperlane* Terbaik (Nugroho, dkk., 2003)

Data *training* dinotasikan dengan $\vec{x}_i \in R^d$ dan *class* dinotasikan dengan $y_i \in \{-1,+1\}$ dimana -1 dan +1 adalah *class* yang berbeda untuk i = 1,2,…n dimana n adalah banyaknya data. Banyak *hyperlane* yang dapat memisahkan data menjadi 2. Berdasarkan metodologi SVM, hanya ada satu *hyperlane* optimal yaitu membentang diantara *maximal margin* (Isa, dkk., 2008). *Hyperlane* terbaik untuk memisahkan *class* didefinisikan dengan persamaan 2.1.

$$\vec{w}.\vec{x} + b = 0 \qquad \ldots(2.1)$$

*Maximal margin* dapat ditemukan dengan mencari titik minimal pada persamaan 2.2 dengan *constraint* pada persamaan 2.3.

$$\min\{1/2\ \|\vec{w}\|^2\} \qquad \ldots(2.2)$$

$$y_i\,(\,\vec{x}_i.\vec{w} + b\,) - 1 \geq 0,\ \forall\, i \qquad \ldots(2.3)$$

UNIVERSITAS MULTIMEDIA NUSANTARA

Persamaan diatas merupakan kasus dimana kedua buah *class* dapat terpisah sempurna. Untuk kasus *class* tidak dapat terpisah dapat dimodifikasi persamaan 2.4 dengan menambahkan *slack variable* $\varepsilon_i$ ($\varepsilon_i > 0$ ).

$$y_i ( \vec{x}_i . \vec{w} + b ) \geq 1 - \varepsilon_i , \forall i \qquad \ldots(2.4)$$

Sehingga persamaan 2.4 menjadi persamaan 2.5.

$$\min\{ 1/2 \|\vec{w}\|^2 + C \sum_{i=1}^{n} \varepsilon_i \} \qquad \ldots(2.5)$$

Paramater C dipilih untuk mengontrol *tradeoff* antara margin dan error klasifikasi $\varepsilon_i$. Nilai C yang besar berarti akan memberikan *penalty* yang lebih besar terhadap error klasifikasi tersebut (Nugroho, dkk., 2003). Nilai C yang kecil akan menghasilkan margin yang lebar namun memberikan *penalty* yang lebih kecil terhadap error klasifikasi. *Minimizing* persamaan 2.5 dengan *constraint* pada persamaan 2.4 menyebabkan *Generalized Optimal Separating Hyperlane*. Persamaan ini merupakan *Quadratic Programming* (QP) *problem* yang dapat diselesaikan dengan metode *Lagrange multipliers*, dengan memaksimalkan *objective function* dari persamaan 2.4 sehingga menjadi persamaan 2.6.

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j \vec{x}_i \vec{x}_j$$

$$\text{dengan } \textit{constraint } \alpha_i \geq 0 \ (i=1,2,\ldots,n), \ \sum_{i=1}^{n} \alpha_i y_i = 0 \qquad \ldots(2.6)$$

Dari hasil perhitungan ini diperoleh $\alpha_i$ yang kebanyakan positif. Data yang berkorelasi dengan $\alpha_i$ positif disebut *support vector* (Nugroho, dkk., 2003).

## 2.8 RBF Kernel

Untuk menyelesaikan *problem* non-linear pada SVM, digunakan *kernel*. Dua jenis *kernel* yang paling sering digunakan adalah *polynomial kernel* dan



Implementasi metode support..., Gandhi Pranata, FTI UMN, 2019

*radial basis function* (*RBF*) *kernel* (Tong & Koller, 2001). *RBF kernel* didefinisikan dengan persamaan 2.7 sebagai berikut.

$$K(x_i,x_j) = \exp(-\gamma \|x_i - x_j\|^2), \forall \gamma > 0 \quad \ldots(2.7)$$

Dimana K( ) adalah fungsi pemetaan kernel, $x_i$ dan $x_j$ adalah ruang input ke ruang fitur, dan $\gamma$ adalah perubahan parameter kernel yang dapat memberikan hasil lebih baik (Prasad, dkk., 2010). Parameter $\gamma$ menentukan seberapa jauh pengaruh dari sebuah *training.* Bila $\gamma$ kecil maka pengaruhnya jauh dan bila $\gamma$ besar berarti pengaruhnya dekat.

## 2.9 K Fold Cross Validation

Pada prosedur *K fold cross validation* setiap model akan dievaluasi sebanyak K kali. Data ditetapkan secara acak dan dibagi menjadi K bagian. Model akan dibangun dari K-1 bagian data yang digunakan sebagai *training set.* Prediksi akurasi dari kandidat model akan dievaluasi pada *test set.* Akurasi didapat dengan dengan proporsi dari total prediksi yang benar (Santra, A.K. & Christy, C.J.). Maka dari itu nilai akurasi dapat dipetakan dengan persamaan 2.8.

$$\text{Akurasi} = \frac{\text{Jumlah prediksi benar}}{\text{Jumlah data keseluruhan}} \quad \ldots(2.8)$$

Secara berurutan menggunakan setiap K bagian sebagai *test set* dan mengulangi pembangunan model dan prosedur pengukuran evaluasi, akan dipilih model terbaik sebagai model optimal (Jung & Hu, 2015).

UNIVERSITAS
MULTIMEDIA
NUSANTARA

## 2.10 Standar Deviasi

Standar deviasi adalah nilai yang digunakan untuk menentukan persebaran data dalam *sample*. Standar deviasi didapat dengan permasaan 2.9 (Geary, R.C.).

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (xi - \mu)^2} \qquad \ldots(2.9)$$

Dimana N merupakan jumlah data, $xi$ merupakan nilai data individu dan $\mu$ merupakan nilai rata-rata. Standar deviasi dengan nilai 0 menunjukkan bahwa seluruh data adalah sama. Nilai standar deviasi yang semakin besar menunjukkan titik data individu semakin jauh dari nilai rata-rata.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA